

## University of Groningen

### Cheryl's birthday

van Ditmarsch, H; Hartley, Michael Ian; Kooi, Barteld; Welton, Jonathan ; Yeo, Joseph B. W.

*Published in:*

Proceedings Sixteenth Conference on <i>Theoretical Aspects of Rationality and Knowledge</i> (TARK 2017)

*DOI:*

[10.4204/EPTCS.251.1](https://doi.org/10.4204/EPTCS.251.1)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*

Final author's version (accepted by publisher, after peer review)

*Publication date:*

2017

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

van Ditmarsch, H., Hartley, M. I., Kooi, B., Welton, J., & Yeo, J. B. W. (2017). Cheryl's birthday. In J. Lang (Ed.), *Proceedings Sixteenth Conference on Theoretical Aspects of Rationality and Knowledge (TARK 2017)* (pp. 1-9). (Electronic Proceedings in Theoretical Computer Science; Vol. 251). Open Publishing Association. <https://doi.org/10.4204/EPTCS.251.1>

#### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

#### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# Cheryl's Birthday

Hans van Ditmarsch

LORIA, CNRS  
University of Lorraine  
Nancy, France

[hans.van-ditmarsch@loria.fr](mailto:hans.van-ditmarsch@loria.fr)

Michael Ian Hartley

DownUnder Geosolutions  
Sydney, Australia  
[mikeh@dug.com](mailto:mikeh@dug.com)

Barteld Kooi

Faculty of Philosophy  
University of Groningen  
Netherlands

[b.p.kooi@rug.nl](mailto:b.p.kooi@rug.nl)

Jonathan Welton

95 Heathermount Drive  
Crowthorne, Berkshire  
United Kingdom

[j\\_welton@hotmail.com](mailto:j_welton@hotmail.com)

Joseph B.W. Yeo

National Institute of Education  
Nanyang Technological University  
Singapore

[josephbw.yeo@nie.edu.sg](mailto:josephbw.yeo@nie.edu.sg)

We present four logic puzzles and after that their solutions. Joseph Yeo designed *Cheryl's Birthday*. Mike Hartley came up with a novel solution for *One Hundred Prisoners and a Light Bulb*. Jonathan Welton designed *A Blind Guess* and *Abby's Birthday*. Hans van Ditmarsch and Barteld Kooi authored the puzzlebook *One Hundred Prisoners and a Light Bulb* [5] that contains other knowledge puzzles, and that can also be found on the webpage <http://personal.us.es/hvd/lightbulb.html> dedicated to the book.

## 1 Logic Puzzles

### 1.1 Cheryl's Birthday

*Albert and Bernard just became friends with Cheryl, and they want to know when her birthday is. Cheryl gives them a list of 10 possible dates for her birthday: May 15, 16, 19; June 17, 18; July 14, 16; and August 14, 15, 17. Cheryl then tells only to Albert the month of her birthday, and tells only to Bernard the day of her birthday. (And Albert and Bernard are aware that she did so.) Albert and Bernard now have the following conversation:*

- *Albert: "I don't know when Cheryl's birthday is, but I know that you don't know either."*
- *Bernard: "At first I didn't know when Cheryl's birthday is, but now I know."*
- *Albert: "Now I also know when Cheryl's birthday is."*

#### When is Cheryl's birthday?

This riddle was set by Joseph Yeo as one of the problems for the Secondary 3 and the Secondary 4 levels of the 2015 Singapore and Asian Schools Math Olympiad (SASMO) contests. He had modified it from a 2013 entry of a blog by 'Mr Brown',<sup>1</sup> that Joseph later traced to a 2006 entry in *The Math Forum 'Ask Dr. Math'*.<sup>2</sup> Cheryl's Birthday riddle is reminiscent of Freudenthal's Sum and Product riddle [6] (see also [5, Chapter 7]), wherein the conversation between Mr. Sum and Mr. Product is very similar.

Cheryl's Birthday riddle subsequently caught a lot of public attention. A Singapore TV presenter, Kenneth Kong, was given this puzzle by a friend whose niece had difficulty solving it. Because the niece was in primary school, Kong thought that the riddle was intended for primary school students. And

---

<sup>1</sup><http://www.mrbrown.com/blog/2013/02/yes-another-psle-maths-question-to-blow-your-mind.html>

<sup>2</sup><http://mathforum.org/library/drmath/view/68613.html>

because Kong and his wife disagreed on the correct answer, he decided to post it online on his Facebook page. This happened on 10 April 2015. Then, as they say, it went viral. It was in subsequent days discussed by Alex Bellos in the British newspaper *The Guardian*, in the BBC news, in the *New York Times*, in *India Today*, etc.<sup>3</sup> See also [9].

## 1.2 One Hundred Prisoners and a Light Bulb

*A group of 100 prisoners, all together in the prison dining area, are told that they will be all put in isolation cells and then will be interrogated one by one in a room containing a light with an on/off switch. The prisoners may communicate with one another by toggling the light-switch (and that is the only way in which they can communicate). The light is initially switched off. There is no fixed order of interrogation, or interval between interrogations, and the same prisoner may be interrogated again at any stage. When interrogated, a prisoner can either do nothing, or toggle the light-switch, or announce that all prisoners have been interrogated. If that announcement is true, the prisoners will (all) be set free, but if it is false, they will all be executed. While still in the dining room, and before the prisoners go to their isolation cells (forever), can the prisoners agree on a protocol that will set them free?*

The 100 prisoners riddle has been going round since the early 2000s, a very nice early elaboration is [3]. Solutions wherein agents play different roles ('counter' and 'follower') are widely known, and also solutions wherein all agents play the same role but wherein part of their behaviour is determined by chance (whether one becomes a 'counter' or a 'follower' is determined by rolling a pair of dice), and wherein termination is not guaranteed but expected (the probability will approach 1). When Hans van Ditmarsch gave a public interest lecture on *One Hundred Prisoners and a Light Bulb* at the University of Western Australia, in Perth, Mike Hartley — he also runs a math puzzle site for children<sup>4</sup> — proposed a solution wherein all agents play the same role (behave the same) and where their behaviour is still determined and not probabilistic. Also in that case, termination is also not guaranteed but expected.

**What uniform protocol can they follow?**

## 1.3 A Blind Guess

*There are three players, each with a card bearing a single digit (0 – 9) on their head such that they can see the other two numbers but not their own — although one is blind and can see no cards. They are told the three digits make up a three digit square number (i.e. between  $10 \cdot 10$  and  $32 \cdot 32$ ), and are asked to call out when they know their own number. After a long, long, long silence, the blind person correctly identifies his card's number.*

**What was it?**

This riddle was proposed by Jonathan Welton. It can be seen as a variation of the *Who has the Sum?* riddle [5, Chapter 6]. That riddle was also originally proposed by him, in the *Sunday Times* magazine Brainteaser column, in the mid 1990s, and was later reprinted as Puzzle 27 in *The Sunday Times Brainteasers* by Victor Bryant [2]. Hans and Barteld did not know this when they wrote their puzzlebook, although they made an effort to trace the origin of the riddles and to report on their earliest known source (their source for this riddle was from 2003). But they found out when Jonathan contacted them, reading about it in the book.

<sup>3</sup>[https://en.wikipedia.org/wiki/Cheryl's\\_Birthday](https://en.wikipedia.org/wiki/Cheryl's_Birthday)

<sup>4</sup><http://www.dr-mikes-math-games-for-kids.com/>

## 1.4 Abby's Birthday

The problem setter announces to two strangers Abby and Barry:

*"You each know the day of the week you were born on, but not the day of the week each other was born on. However, I can tell you that you were born a day apart, and that Barry was not born on a Monday."*

He then proceeds to ask Abby and Barry alternately if they can deduce which day of the week the other person was born on, and receives the answers:

- Abby: "No."
- Barry: "No."
- Abby: "No."
- Barry: "No."
- Abby: "No."

**On what day of the week was Abby born?**

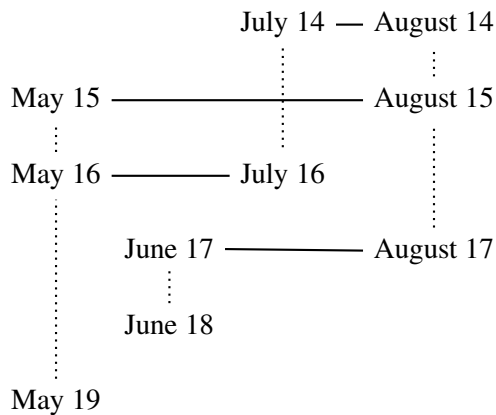
Abby's Birthday problem was also proposed by Jonathan Welton.

## 2 Solutions

We now proceed with the solutions of the riddles. It is more fun if you first try to solve them yourself.

### 2.1 Cheryl's Birthday

We can visually represent the situation wherein Albert and Bernard have just been told their private information about the month and the day, respectively, as follows.

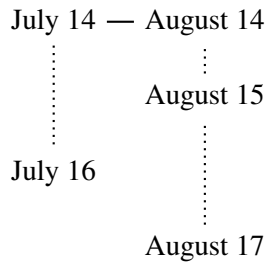


Dates that share the same month are connected by dotted lines. These indicate that Albert cannot distinguish these. When two dates have the same day, they are connected by a solid line, indicating that Bernard cannot distinguish them.

When you look at this picture, June 18th and May 19th are special in the sense that they are not connected to other dates by solid lines. That means that if one of those dates were Cheryl's birthday, then Bernard would know her birthday, because there are no alternatives for him to consider.

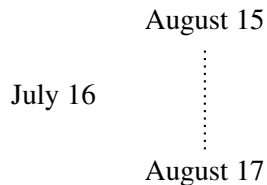
Note also that in this picture for any date there is another date connected to it by a dotted line. This means that for Albert, there is always an alternative. So, whatever Cheryl's birthday is, Albert won't know when. This also means that when he says that he does not know her birthday, he does not convey any information. This was common knowledge even to an outsider. However, when Albert says that he knows that Bernard does not know Cheryl's birthday, we can infer that Cheryl's birthday is not in May or June, because if it were in May or June — as far as Albert is concerned — her birthday might have been on the 19th or 18th and then Bernard would know. The only way that Albert can know that Bernard does not know the birthday, is if this is not the case.

If we update the model with the information that Cheryl's birthday is not in May or June we find the following picture.



In this picture the only dates connected by a solid lines are July 14th and August 14. Consequently, on those dates Bernard does not know when Cheryl's birthday is, but on all of the other dates he would know, since in the updated picture there are no alternative dates for him to consider. (Not only did we learn from Albert's assertion that May and June are excluded, Bernard learnt the same.) So when Bernard states that he does now know when Cheryl's birthday is, we can conclude her birthday is not on the 14th. (We can disregard his statement that at first he did not know Cheryl's birthday, because Albert already asserted that.)

If we carry out the update we get the following picture.



In this picture we see that Albert is still not certain of Cheryl's birthday if it is in August, but otherwise he knows her birthday, because there are no alternatives. He states that he knows, and therefore we now also know that Cheryl's birthday is on:

July 16

Joseph Yeo ventures that *Cheryl's Birthday* was so widely discussed in the media because there was much discussion whether the answer should be July 16 or August 17. In [9, page 4] he explains in detail how some people obtained the wrong answer August 17.

## 2.2 One Hundred Prisoners and a Light Bulb

A well-known solution consists of the prisoners appointing one of them as the *counter*, where the remaining prisoners are *followers*. A follower turns the light on when it is off, but does so only once. If the

light is on he does nothing and if the light is off and he has turned it on before, he also does nothing. The counter turns the light off when it is on, and keeps count of how many times he has turned the light off; and does nothing when the light is on. When the counter has reached 100 he announces that everybody has been interrogated. In this standard solution we can think of all the prisoners as initially having a token with value 1 in their head. The followers will drop one of their token when they turn on the light and remain with value 0. Whereas the counter, by turning off the light, only increases his token all the time, until it has value 100.

There is a probabilistic solution where all prisoners can both act as the counter and as a follower but do so depending on the value of their token and depending on the outcome of rolling a dice: given the value of their token and the state of the light they may *either*,

when the light is off, turn on the light and reduce their token, and, when the light is on, do nothing (thus acting as a follower);

*or*,

when the light is off, do nothing and when the light is on, turn the light off and increase their token (thus acting as a counter).

Termination cannot be guaranteed but can be expected (the probability will approach 1 in due time). Expected execution time (given random scheduling of interrogations) can be shortened if the probability of being a counter is high (or even 1) once a prisoner holds a token with value more than 50 (although this has not been tested by simulations or verified by calculations). Roles are not fixed: for token values below 50 it is important that the probability of being counter or follower is not 0 and not 1. For example, if there are two prisoners hold tokens of value 50, and they only act as counters for that value of token, neither of them will ever turn on the light. Nobody else will anyway. So then the interrogations will go on forever without termination. No prisoner will reach token value 100. For more details on this solution, see [5, page 92, Protocol 6].

It came therefore as a surprise to Hans that Mike Hartley suggested a solution wherein all prisoners play the same role (and wherein that role is not probabilistic), i.e., wherein the state of the light and the value of their token determine whether they act as counter or as follower. It is as follows.

- Each prisoner initially holds a token valued 1.
- If a prisoner enters the room and the light is off,
  - then if their token is 0, they do nothing;
  - otherwise, they switch the light on and subtract 1 from their token.
- If a prisoner enters the room and the light is on,
  - then if their token is 0, they do nothing;
  - and if their token is 99, they announce that all prisoners have been interrogated;
  - otherwise, they switch off the light and add 1 to their token.

Given random scheduling this protocol is again expected to terminate. This may be somewhat surprising, as ‘on average’ a prisoner will turn the light off equally often as turning the light on. The reason is the (known) deviation from this expected average. A large deviation may result in a prisoner substantially increasing his token, for a while, after which it may just as well substantially decrease again. Once this deviation hits the roof, token value 100, the protocol terminates. The longer the interrogations, the more likely it becomes to hit the roof. Clearly, this may take a while.

Now for the standard protocol, given uniform random scheduling, and on assumption a single interrogation takes place every day, the prisoners can expect to get out of prison after 24,5 years. For the prisoners, who of course only care to get out of prison as soon as possible, this outlook is already not quite so rosy. But this novel protocol is much worse. Mike Hartley:

*I ran a computer simulation of it just now, and the 100 prisoners got out in, on average, 1613 years. The luckiest group got out in 'only' 137 years, the unluckiest were imprisoned when the ancient Sumerians were just getting started, and still have 1800 years of their sentences to serve.*

Suppose  $k$  out of  $n$  prisoners have token values  $a_1, a_2, \dots, a_k$ , the rest having token value 0. Let the expected time for the prisoners to escape be  $U_{a_1, a_2, \dots, a_k}$  if the light is off, and  $T_{a_1, a_2, \dots, a_k}$  if the light is on. Note that  $U_n = 0$ , since in that case the prisoners have already escaped. If a random prisoner is chosen to be interrogated, then

$$U_{a_1, a_2, \dots, a_k} = 1 + \frac{n-k}{n} U_{a_1, a_2, \dots, a_k} + \frac{1}{n} \sum_{i=1}^k T_{a_1 - \delta_{1i}, a_2 - \delta_{2i}, \dots, a_k - \delta_{ki}}$$

and

$$T_{a_1, a_2, \dots, a_k} = 1 + \frac{n-k}{n} T_{a_1, a_2, \dots, a_k} + \frac{1}{n} \sum_{i=1}^k U_{a_1 + \delta_{1i}, a_2 + \delta_{2i}, \dots, a_k + \delta_{ki}}.$$

This gives a system of linear equations for the expected times. It is hard to find general formulae for  $U_{a_1, a_2, \dots, a_k}$  and  $T_{a_1, a_2, \dots, a_k}$ , but it is clear that the solution is finite, and so the prisoners will not be imprisoned forever.

For example, to show an endgame, consider we have reached in the 100 prisoners case that one prisoner has a token worth 99, and that the light is on. We now get that  $T_{99} = 1 + \frac{99}{100} T_{99} + \frac{1}{100} U_{(99+1)}$ . Given that  $U_{100} = 0$ ,  $T_{99}$  must therefore be 100. What does this mean? Given a prisoner with token 99, the light on (token 1) and all other 99 prisoners with token 0, there is a  $\frac{1}{100}$  chance that the token-99 prisoner is interrogated and collects the last token, and there is a  $\frac{99}{100}$  chance that a token-0 prisoner is interrogated, in which case the light remains on (back to  $T_{99}$ ). The probability of  $\frac{1}{100}$  indeed comes with an expectation of 100 days. The expected time to escape for  $U_{100}$  is 0 because the prisoner with token 100 declares victory: all prisoners have been interrogated.

For three prisoners we get that  $T_{a,b,c}$  equals:

$$\left(1 + \frac{1}{a+b+c}\right) \cdot [2(a^2b + b^2c + c^2a + ab^2 + bc^2 + ca^2) + 3abc + a + b + c]$$

so that  $T_{1,1,1} = \left(1 + \frac{1}{3}\right) [12 + 3 + 3] = 24$ . To have to wait 24 days to get out of prison is not so bad.

### 2.3 A Blind Guess

There are 22 square numbers between 100 and 999. Every agent carries a number that is one of the three digits of such a square number. A model of the uncertainty about the squares therefore consists triples of digits. Now these are again, in a way, numbers. However, they are numbers between 1 and 999 — where 1 is in fact 001. So, to model this problem we need far more than 22 states. Let the agents be  $A, B, C$  and let  $C$  be the blind agent. We only need to model the uncertainty of  $A$  and  $B$ . Blind agent  $C$  plays the role of us, the problem solvers. The domain is partitioned into a great deal of parts that are  $\{A, B\}$ -connected and that are all disconnected from each other. Let us consider some such parts.

Agent  $A$  cannot distinguish the states connected by a solid line and agent  $B$  cannot distinguish the states connected by a dotted line.

100 — 400 — 900

Here,  $A$  cannot distinguish three states wherein she sees that  $B$  and  $C$  both have 0. She is uncertain whether she has 1, 4, or 9. But  $B$  can distinguish all states. So, when asked to call out if he knows his number, he will say: “Yes.” But the problem description tells us he did not. We can therefore rule out the states 100, 400 and 900 from the model of uncertainty. We cannot *yet* rule out those squares. For that, we also need to consider the part of the model

010 ..... 040 ..... 090

in which case  $A$  will say that she knows her number, and the three  $\{A, B\}$ -disconnected parts of the model

001      004      009

in which case both  $A$  and  $B$  will call out that they know their number.

More interesting cases are, for example,

526 ..... 576 — 676

and

675 ..... 625 — 225 ..... 265 — 765

Note that the squares 625 and 576 occur in these parts of the model in all their permutations. Still, the parts do not have any symmetry.

If we take the first part, then either  $A$  or  $B$  would have initially called out (namely if the state had been, respectively, 526 or 676), or else both would have called out later. But as this did not happen, after two updates this part of the model disappears.

Similarly for the other part, after three updates. Agent *C* then reasons at follows: the state cannot be 675 because then *A* would have called out. But it cannot be 625 either because then, after *B* does not hear *A* call out, *B* would have called out. And it cannot be 225 either, because, after *A* does not hear *B* call out because *B* did not hear *A* call out, *A* would have called out. And dually for *B*, from the other side of this  $\{A, B\}$ -chain.

We did not depict various other parts of the domain. After a certain finite number of such updates, there is no further uncertainty reduction. The domain will then be reduced to the following, unique, circular chain of uncertainty.

$$\begin{array}{ccccc} 484 & \text{---} & 184 & \text{.....} & 144 \\ \vdots & & & & | \\ 414 & \text{---} & 814 & \text{.....} & 844 \end{array}$$

We observe that  $C$  holds 4 in all these states. This is therefore the solution of the problem.

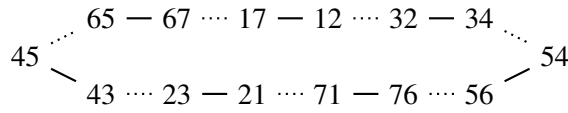
We could actually think of  $C$  as having the universal relation on the domain — he is unable to distinguish any of the states. Successive absence of announcements of knowledge will shrink the domain more and more until at some stage only the circular partition remains. In this we have taken some ‘mathematical licence’ interpreting the problem formulation. There is no explicit synchronization in the ‘after a long, long silence’ part of the description. (By the way, this is the case in all older formulations of such knowledge and ignorance puzzles, such as Littlewood’s 1953 *Miscellany* [7], and prior to that).



We could alternatively have formulated it more explicitly: the problem setter asks  $A$  and  $B$  to give a shout when he claps his hands, and tells them that he will continue to repeat this clapping action until ... Until nothing happens? Then you can wait a long time! We would then have to specify the number of repetitions of hand clapping, like 3, or 4 (3 is in fact enough). We find the original formulation then more elegant, as the 'long, long silence' suggests a *finite* number of iterations, without specifying the amount. Which is all the information we need as problem solvers.

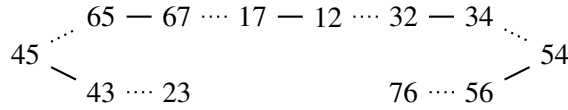
## 2.4 Abby's Birthday

We number the days of the week from 1 to 7 for Monday to Sunday. Let a pair  $ij$  stand for 'Abby was born on a  $i$  and Barry was born on a  $j$ '. The model of uncertainty then is as follows, where Abby is uncertain between worlds linked by a solid line and Barry is uncertain between states linked by a dotted line.

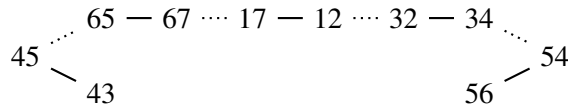


The successive announcements result in the following uncertainty reductions.

- Problem setter: "Barry was not born on Monday."



- Abby: "No."



- Barry: "No."
- Abby: "No."
- Barry: "No."
- Abby: "No."

17 — 12

Therefore, Abby was born on Monday!

## 3 Dynamic Epistemic Logic

Riddles about knowledge and ignorance such as those in this contribution have stood at the birth of the so-called *dynamic epistemic logic*, that is a modal logic of knowledge and change of knowledge. The models visualized are known as *Kripke models*. A well-known dynamic epistemic logic is *public announcement logic*. For example, in Cheryl's Birthday problem, Albert saying 'I know that you don't know' can be formalized in the logic of knowledge as the formula  $K_A \neg \bigvee_{p_d} K_B p_d$ , where  $K_A$  stands for

‘Albert knows that’ and  $K_B$  for ‘Bernard knows that’,  $\neg$  is logical negation,  $\vee$  stands for disjunction, and where  $p_d$  is an atom ranging over the possible birthdays  $d$ . In the given model, this formula is true exactly on dates in July or August and false in dates in May and June. Albert’s so-called public announcement of this formula therefore results in the model that is restricted to dates in July and August. Bernard subsequently saying ‘Now I know’ corresponds to updating with (a public announcement of) the formula  $\bigvee_{p_d} K_B p_d$ , and Albert’s final announcement to an update with  $\bigvee_{p_d} K_A p_d$ . For more information on dynamic epistemic logic see for example [4, 8, 1].

## References

- [1] A. Baltag & B. Renne (2016): *Dynamic Epistemic Logic*. In E.N. Zalta, editor: *The Stanford Encyclopedia of Philosophy*.
- [2] V. Bryant (2002): *The Sunday Times Brainteasers*. Times Books.
- [3] P. Dehay, D. Ford & H. Segerman (2003): *One hundred prisoners and a lightbulb*. *Mathematical Intelligencer* 25(4), pp. 53–61, doi:10.1007/BF02984862.
- [4] H. van Ditmarsch, W. van der Hoek & B. Kooi (2005): *Playing Cards with Hintikka: An introduction to dynamic epistemic logic*. *The Australasian Journal of Logic* 3, pp. 108–134. Available at <http://www.philosophy.unimelb.edu.au/ajl/2005/>.
- [5] H. van Ditmarsch & B. Kooi (2015): *One Hundred Prisoners and a Light Bulb*. Copernicus, doi:10.1007/978-3-319-16694-0.
- [6] H. Freudenthal (1969): *Problem No. 223 (formulation of the Sum and Product problem)*. *Nieuw Archief voor Wiskunde* 3(17), p. 152.
- [7] J.E. Littlewood (1953): *A Mathematician’s Miscellany*. Methuen and Company.
- [8] L. Moss (2015): *Dynamic Epistemic Logic*. In H. van Ditmarsch, J. Halpern, W. van der Hoek & B. Kooi, editors: *Handbook of epistemic logic*, College Publications, pp. 261–312.
- [9] J.B.W. Yeo (2016): *Tracing the origin of Cheryl’s birthday puzzle*. *Maths Buzz* 16(1), pp. 2–4.